

AMENDMENTS TO THE CLAIMS

Please amend Claims 1, 8, 21, and 28 as follows:

1. (Currently Amended) A method, comprising:

receiving a request from a user for a web page at a first web address, the first web address including a hostname;

determining traffic loads of a plurality of mirrored customer web servers among a customer's plurality of web servers, each of the customer web servers storing the web page;

determining a customer web server from the plurality of mirrored customer web servers that is appropriate for the request, the customer web server having a traffic load lower than traffic loads of remaining customer web servers from the plurality of mirrored customer web servers;

determining an IP address of the customer web server;

directing the request from the user to the customer web server;

receiving a request from the user for static content on the web page at a second web address, the second web address including the hostname;

determining service metrics of caching servers in a network of caching servers different from the customer's plurality of web servers;

wherein a customer pays a fee to a service for use of the network of caching servers storing static content for the customer;

determining a caching server from the network of caching servers that is appropriate for the request for static content, the caching server having service metrics better than service metrics of remaining caching servers from the network of caching servers;

retrieving the static content from the caching server; and

providing the static content to the user.

2. (Previously Presented) The method of claim 1 further comprising:

determining load of caching servers in the network of caching servers;

wherein the determining the caching server from the network of caching servers that is appropriate for the request step selects a caching server having a latency and a load lower than latency or load of remaining caching servers from the network of caching servers.

3. (Previously Presented) The method of claim 1 further comprising:
determining whether the caching server includes the static content;
determining a customer web server that includes the static content when the caching server does not include the static content;
retrieving the static content from the web server that includes the static content; and
storing the static content from the web server in the caching server.

4. (Previously Presented) The method of claim 3 wherein the determining the web server step comprises:
determining traffic loads of the plurality of mirrored customer web servers, each of the customer web servers storing the static content; and
determining a second customer web server from the plurality of mirrored customer web servers that is appropriate for the request, the second customer web server having a traffic load lower than traffic loads of remaining customer web servers from the plurality of mirrored customer web servers.

5. (Previously Presented) The method of claim 4 wherein retrieving the static content from the web server step comprises:
determining an IP address of the second customer web server; and
requesting the static content from the second customer web server at the second customer web server IP address.

6. (Previously Presented ded) The method of claim 1 wherein the network of caching servers includes a domain name server.

7. (Previously Presented) The method of claim 1 wherein the request from the user for the web page is transferred from a first domain name server;
wherein the network of caching servers includes a second domain name server; and
wherein the second domain name server determines the customer web server from the plurality of mirrored customer web servers.

8. (Currently Amended) A method, comprising:
receiving a first request from a client DNS server to resolve a first domain name, the client DNS server receiving a request from a user of a web page address that includes the first domain name;
determining load measurements of a plurality of mirrored customer web servers among a customer's plurality of web servers, each of the customer web servers addressable by the first domain name, and each of the customer web servers configured to service the request from the user;
determining a customer web server from the plurality of mirrored customer web servers, the customer web server having a traffic load lower than traffic loads of other customer web servers from the plurality of mirrored customer web servers;
determining an IP address of the customer web server;
providing the IP address of the customer web server to the client DNS server;
receiving a second request from the client DNS server to resolve a second domain name, the client DNS server receiving a request from the user of a uniform resource locator that includes the second domain name;

determining performance metric measurement of caching servers in a network of caching servers different from the customer's plurality of web servers, each of the caching servers addressable by the second domain name;

wherein a customer pays a fee to a service for use of the network of caching servers storing static content for the customer;

determining a caching server from the network of caching servers, the caching server having performance metrics lower than performance metrics of other caching servers from the network of caching servers;

providing an IP address of the caching server to the client DNS server;

retrieving data from the caching server in response to the uniform resource locator; and

providing the data to the user.

9. (Original) The method of claim 8 wherein the load measurements comprise latency measurements.

10. (Previously Presented) The method of claim 8 wherein the performance metric measurements comprise any of: load CPU and memory measurements, HTTP response measurements, and FTP response measurements.

11. (Previously Presented) The method of claim 8 wherein retrieving data from the caching server step comprises:

determining whether the caching server includes the data;

retrieving data from a second customer web server from the mirrored customer web servers when the caching server does not include the data; and

storing the data within the caching server.

12. (Previously Presented) The method of claim 11 wherein retrieving data from the second customer web server step comprises:

determining the second customer web server from the plurality of mirrored customer web servers, the second customer web server having a traffic load lower than traffic loads of remaining customer web servers from the plurality of mirrored customer web servers; and

retrieving the data from the second customer web server.

13. (Previously Presented) The method of claim 8 further comprising: receiving a first request from a second client DNS server to resolve a third domain name, the second client DNS server receiving a request from a second user of a second web page address that includes the third domain name;

determining load measurements of a plurality of second mirrored customer web servers among a customer's plurality of web servers, each of the second customer web servers addressable by the third domain name, and each of the second customer web servers storing data configured to service the request from the second user;

determining a second customer web server from the plurality of second mirrored customer web servers, the second customer web server having a traffic load lower than traffic loads of other second customer web servers from the plurality of second mirrored customer web servers;

determining an IP address of the second customer web server; and providing the IP address of the second customer web server to the second client DNS server.

14. (Previously Presented) The method of claim 13 further comprising:

receiving a second request from the second client DNS server to resolve the second domain name, the second client DNS server receiving a request from the second user of a second uniform resource locator that includes the second domain name;

retrieving a second set of data from the caching server in response to the second uniform resource locator; and
providing the second set of data to the user.

15-20. (Canceled)

21. (Currently Amended) An apparatus, comprising:

a module for receiving a request from a user for a web page at a first web address, the first web address including a hostname;

a module for determining traffic loads of a plurality of mirrored customer web servers among a customer's plurality of web servers, each of the customer web servers storing the web page;

a module for determining a customer web server from the plurality of mirrored customer web servers that is appropriate for the request, the customer web server having a traffic load lower than traffic loads of remaining customer web servers from the plurality of mirrored customer web servers;

a module for determining an IP address of the customer web server;

a module for directing the request from the user to the customer web server;

a module for receiving a request from the user for static content on the web page at a second web address, the second web address including the hostname;

a module for determining service metrics of caching servers in a network of caching servers different from the customer's plurality of web servers;

wherein a customer pays a fee to a service for use of the network of caching servers storing static content for the customer;

a module for determining a caching server from the network of caching servers that is appropriate for the request for static content, the caching server having service metrics better than service metrics of remaining caching servers from the network of caching servers;

a module for retrieving the static content from the caching server; and

a module for providing the static content to the user.

22. (Previously Presented) The apparatus of claim 21 further comprising:

a module for determining load of caching servers in the network of caching servers;

wherein determining the caching server from the network of caching servers that is appropriate for the request module selects a caching server having a latency and a load lower than latency or load of remaining caching servers from the network of caching servers.

23. (Previously Presented) The apparatus of claim 21 further comprising:

a module for determining whether the caching server includes the static content;

a module for determining a customer web server that includes the static content when the caching server does not include the static content;

a module for retrieving the static content from the web server that includes the static content; and

a module for storing the static content from the web server in the caching server.

24. (Previously Presented) The apparatus of claim 23 wherein the determining the web server module comprises:

a module for determining traffic loads of the plurality of mirrored customer web servers, each of the customer web servers storing the static content; and

a module for determining a second customer web server from the plurality of mirrored customer web servers that is appropriate for the request, the second customer web server having a traffic load lower than traffic loads of remaining customer web servers from the plurality of mirrored customer web servers.

25. (Previously Presented) The apparatus of claim 24 wherein the retrieving the static content from the web server module comprises:

a module for determining an IP address of the second customer web server; and

a module for requesting the static content from the second customer web server at the second customer web server IP address.

26. (Previously Presented) The apparatus of claim 21 wherein the network of caching servers includes a domain name server.

27. (Previously Presented) The apparatus of claim 21 wherein the request from the user for the web page is transferred from a first domain name server;

wherein the network of caching servers includes a second domain name server; and

wherein the second domain name server determines the customer web server from the plurality of mirrored customer web servers.

28. (Currently Amended) An apparatus, comprising:
a module for receiving a first request from a client DNS server to resolve a first domain name, the client DNS server receiving a request from a user of a web page address that includes the first domain name;

a module for determining load measurements of a plurality of mirrored customer web servers among a customer's plurality of web servers, each of the customer web servers addressable by the first domain name, and each of the customer web servers configured to service the request from the user;

a module for determining a customer web server from the plurality of mirrored customer web servers, the customer web server having a traffic load lower than traffic loads of other customer web servers from the plurality of mirrored customer web servers;

a module for determining an IP address of the customer web server;
a module for providing the IP address of the customer web server to the client DNS server;

a module for receiving a second request from the client DNS server to resolve a second domain name, the client DNS server receiving a request from the user of a uniform resource locator that includes the second domain name;

a module for determining performance metric measurement of caching servers in a network of caching servers different from the customer's plurality of web servers, each of the caching servers addressable by the second domain name;

wherein a customer pays a fee to a service for use of the network of caching servers storing static content for the customer;

a module for determining a caching server from the network of caching servers, the caching server having performance metrics lower than performance metrics of other caching servers from the network of caching servers;

a module for providing an IP address of the caching server to the client DNS server;

a module for retrieving data from the caching server in response to the uniform resource locator; and

a module for providing the data to the user.

29. (Previously Presented) The apparatus of claim 28 wherein the load measurements comprise latency measurements.

30. (Previously Presented) The apparatus of claim 28 wherein the performance metric measurements comprise any of: load CPU and memory measurements, HTTP response measurements, and FTP response measurements.

31. (Previously Presented) The apparatus of claim 28 wherein retrieving data from the caching server comprises:

a module for determining whether the caching server includes the data;
a module for retrieving data from a second customer web server from the plurality of mirrored customer web servers when the caching server does not include the data; and

a module for storing the data within the caching server.

32. (Previously Presented) The apparatus of claim 31 wherein the retrieving data from the second customer web server module comprises:

a module for determining the second customer web server from the plurality of mirrored customer web servers, the second customer web server having a traffic load lower than traffic loads of remaining customer web servers from the plurality of mirrored customer web servers; and

a module for retrieving the data from the second customer web server.

33. (Previously Presented) The apparatus of claim 28 further comprising:

a module for receiving a first request from a second client DNS server to resolve a third domain name, the second client DNS server receiving a request from a second user of a second web page address that includes the third domain name;

a module for determining load measurements of a plurality of second mirrored customer web servers among a customer's plurality of web servers, each of the second customer web servers addressable by the third domain name, and each of the second customer web servers storing data configured to service the request from the second user;

a module for determining a second customer web server from the plurality of second mirrored customer web servers, the second customer web server having a traffic load lower than traffic loads of other second customer web servers from the plurality of second mirrored customer web servers;

a module for determining an IP address of the second customer web server; and

a module for providing the IP address of the second customer web server to the second client DNS server.

34. (Previously Presented) The apparatus of claim 33 further comprising:

a module for receiving a second request from the second client DNS server to resolve the second domain name, the second client DNS server receiving a request from the second user of a second uniform resource locator that includes the second domain name;

a module for retrieving a second set of data from the caching server in response to the second uniform resource locator; and

a module for providing the second set of data to the user.